An Empirical Investigation of Google Trends for Assessing Investor's Sentiments and Predictability of Indian Stock Market Returns

SHORT SYNOPSIS

Submitted in partial fulfillment of the requirement of the degree of

DOCTOR OF PHILOSOPHY

to

Manav Rachna International Institute of Research and Studies

(Deemed to be University)

by

DIVYA JAIN
19/Ph.D/054

Under the supervision of

Supervisor
Dr. Meghna Chhabra
Associate Professor, Faculty of Management Studies

FACULTY OF MANAGEMENT STUDIES

MANAV RACHNA INTERNATIONAL INSTITUTE OF RESEARCH & STUDIES

SECTOR-43, SURAJKUND –DELHI ROAD, FARIDABAD – 121001

FACULTY OF MANAGEMENT STUDIES

Certificate

I hereby declare that the short synopsis **"An Empirical Investigation of Google Trends for Assessing Investor's Sentiments and Predictability of Indian Stock Market Returns"** being submitted in partial fulfilment of the requirement for the degree of Doctor of Philosophy in **Finance** under **Faculty of Management Studies** of Manav Rachna International Institute of Research & Studies, Faridabad during the academic year 2020-21. It is a bonafide record of original work carried out under the supervision of '**Dr. Meghna Chhabra** has not been submitted in full or part to any other university or institute for a degree or diploma.
Further, it is certified that the short synopsis is plagiarism-free and the proof has been submitted.

Divya Jain
19/Ph.D/054

6/10/2021

Signature of the Supervisor

06/10/2021

(Chairperson, DRC)

(Chairperson, CDP)

ABSTRACT

The evolving research in the field of behavioral finance puts a question on traditional finance theories and incorporates psychology and social sciences along with the mathematical component. The human emotions and behavior of investor do affect the market movement, and the traditional finance cannot fully explain this. Over the years, due to the absence of standard measures to enumerate investor sentiment, various ways have been adopted to gauge the same. To determine current trends and opinions, web search is a rich source of real-time information. The public mood reflected on the internet reflect the whole of society on a large scale. There is evident penetration of the internet into human lives, and this has led to growth in interest of researchers in the area. This study proposes one such attempt to explore the investor's sentiments represented through their web activities and its influence on stock market movements. The current study proposes to explore the forecasting power of investors' sentiment on returns. It focuses on determining that how in the light of behavioral finance, the Google Search Volume Index (GSVI) can be proxied for investor's sentiments while taking financial investment decisions and its utility to improve the predictability of stock market returns.

Keywords: Behavioral Finance, Investors Sentiment, Google Search Volume Index (GSVI), Predicting Stock Market Returns

CONTENTS

Description	Page No
Introduction	1-7
Literature Review	8-12
Research Gaps	13
Description of the Topic	14
Objectives of the Study	15
Proposed Research Methodology	16-18
Proposed Outcome of the Research	19
References	
Plagiarism Report	

ABBREVIATION INDEX

Abbreviation	Description
BSE	Bombay Stock Exchange
CCI	Consumer Confidence Index
CSI	Consumer Sentiment Index
EMH	Efficient Market Hypothesis
EMT	Efficient Market Theory
GIS	Google Insights for Search
GSVI	Google Search Volume Index
II Index	Investor Intelligence Index
MAPE	Mean Absolute Percentage Error
NSE	National Stock Exchange
US	United States

INTRODUCTION

Traditional finance based on models, principles, and theories states that markets are efficient and systematic. It is grounded on the assumption of rational behavior of investors and states that estimation of prices of financial assets incorporates all the possible available information and is termed as EMH (Efficient Market Hypothesis) (Kumar & Goyal, 2015). Sometimes the underlying fundamentals do not justify the extreme market fluctuations and price movements. There comes in the concept of Behavioral Finance, which justifies the role of investor's behavior. The behavior of investors encompasses a psychological and social sciences approach to the financial investments decision-making process. Both traditional and behavioral finance have different schools of thought behind the shaping of stock market movements but share a common belief that the volatility and market movements are shaped based on the absorption of information. With the passage of time, the role of investor's sentiments is justified by research on the theoretical framework. The real question is no longer related to the determination of whether the stock prices are affected by sentiments or not, but rather quantification of its effect and means by which investor sentiments can be measured, (Baker & Wurgler, 2007).

Behavioral Finance and Investor Sentiments

The question of the efficiency of markets due to increasing anomalies has revolutionized the study of financial theories. The influence of the behavior of an investor in his financial decision-making cannot be ignored and must be given due importance. Economists have identified that the extreme fluctuations in the market cannot be explained only by the underlying fundamentals of the financial markets. Literature explains various psychological concepts and numerous theories which are applicable in the field of behavioral finance. The research work from Paule-Vianez, Gómez-Martínez & Prado-Román (2020) conducted a bibliometric study in the field of behavioural finance. The work analyses the domain not just as a whole subject but also identifies different themes underlying in this umbrella term of behavioral finance. By performing a co-word analysis, 13 different themes within behavioral finance have been identified, and the level of development under each theme is assessed. Figure 1 is retrieved from the work done by Paule-Vianez et al. (2020) and the figure positions themes like investor sentiments, expected stock return, disposition effect and overconfidence as motor themes. The theme which remarkably stands out in its influence and impact is investor sentiment. It has been detected that from the perspective of modern finance, the study of the relationship of investor sentiment on the decision-making process could be one of the significant research area in the field of finance.

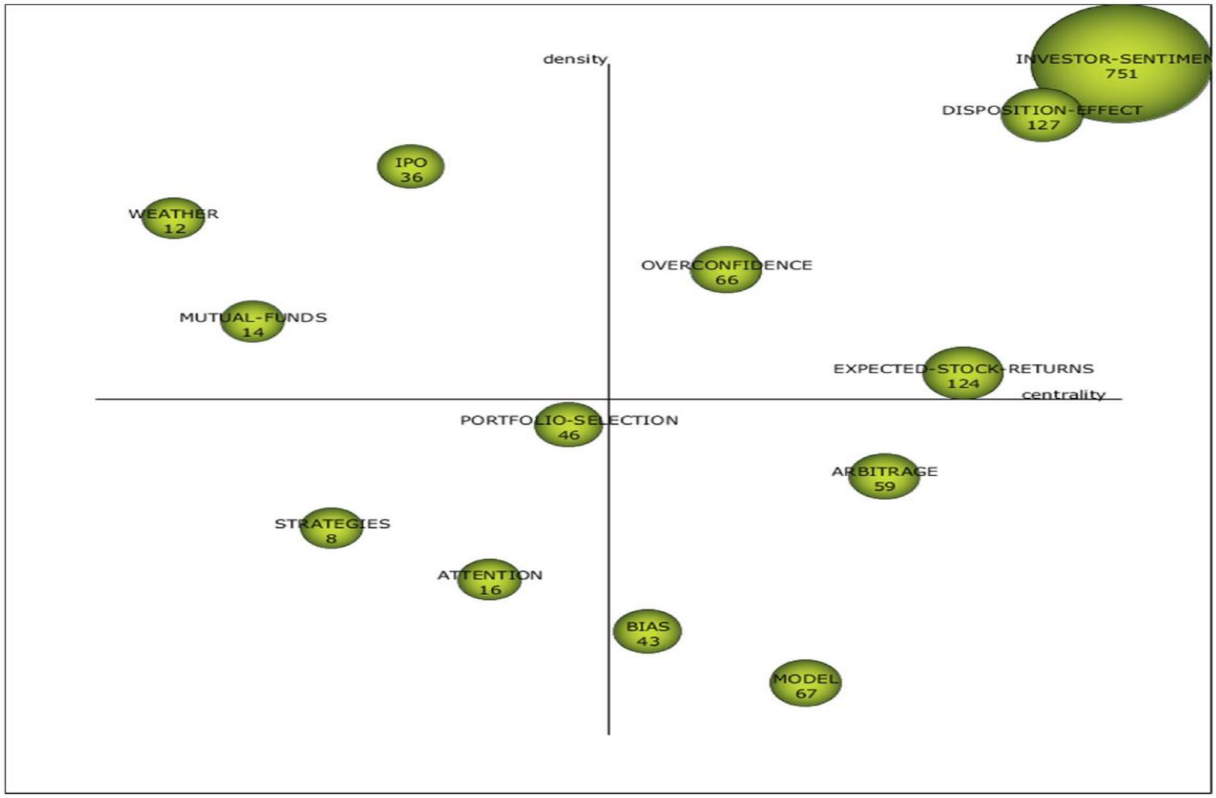

Figure 1. 13 Different themes under Behavioral Finance (defined by Paule-Vianez et al., 2020)

The investor sentiment has emerged as a prolific field of research. Sentiment is the field that stands out. (Baker & Wurgler, 2007) defined investor sentiments as "belief about future cash flows and investment risks that is not justified by the facts at hand". Sentiment is basically the perception of an individual which directly influences his decision-making, and the same is reflected in financial markets. So, investor sentiment can be considered as a relevant factor to explain the investors' behavior in financial markets. Since no universal standard measure is there to determine investor sentiments. To measure sentiment, various methods like surveys, indexes based on macro-economic variables have been utilized in past years. In recent years, new ways of measuring public mood and sentiment have emerged in the form of social media posts, blogs, news headlines, or web search-based data.

Google as a Web Search Engine

The rapid evolution of the digital landscape has made people spend considerably more time on the internet using connected technology. A report by Digital Market Outlook on Statista.com (2020) states that in 2020, the number of internet users in India has increased to 696.77 million from 302.36 million in 2015 and it is expected to grow to a whopping number of 974.86 million

by 2025. This continuous increase in internet users posits the need to examine the information accumulated through internet sources.

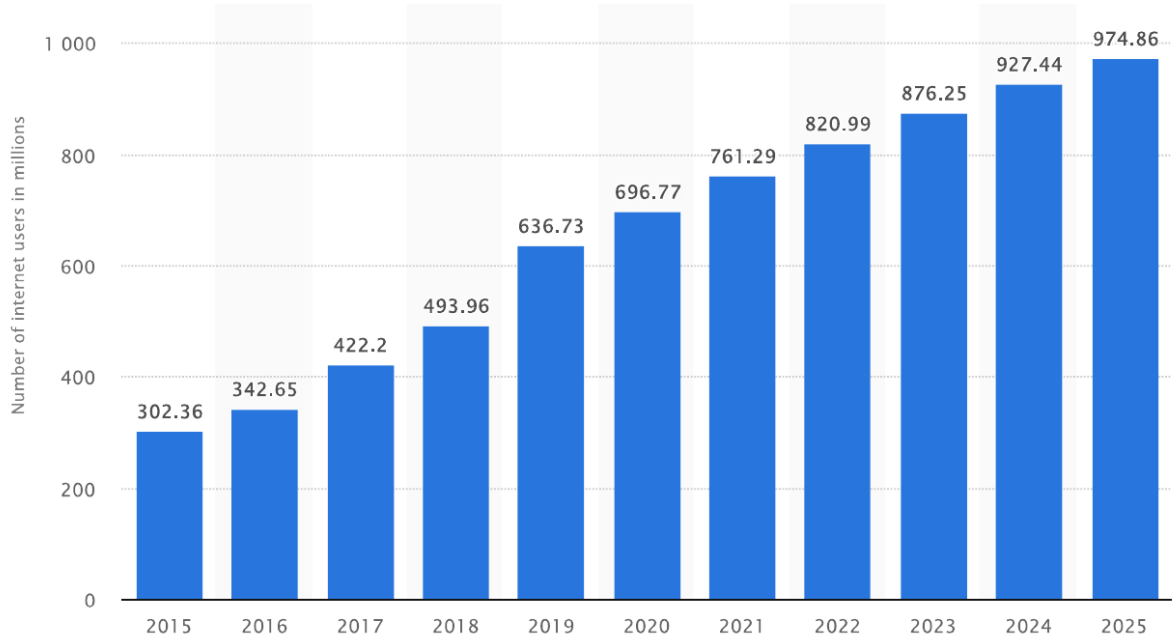

Figure 2. Number of Internet Users in India, Digital Market Outlook (Statista, 2020)

Google as a search engine is one of the extremely utilized platforms to gather information across the globe and a very high percentage of 90% of internet searches are handled by the giant (Harford, 2017). Google is a powerhouse and most sought after go to the search engine with approximately 90% of worldwide search engines market share. Google receives over 71,000 searches per second on any given day and is arguably the most trusted source of information out there. The following table represents the percentage share of search engines in the world's search traffic in the year 2018-2019.

Table 1. Percentage Share of Search Engines in World's Search Traffic (2018-2019)

Search Engine	**Percentage Share in World's Search Traffic**
Google	77%
Bing	9.41%
Baidu	9.34%
Yahoo	3%
Yandex	1.23%
Others	0.02%

(Source: Author's own compilation)

The search volume difference between Google and the other search engines is wide enough to signify the Google global leadership. Google is the most popular search engine in India as well as many countries across the globe.

Google Web Search Data and Field of Research

Undoubtedly, the web has evolved as a framework which involves a lot of information dissemination among its users. Due to technological advances, the broad communication channels, including social media platforms like blogs, microblogging sites, social news, web search data, etc., plays an important role in processing timely information to the masses. The popularity of social media is not unknown to anyone, and the data gathered from the web search and social media is the collective representative of thoughts and ideas about several aspects of the world. About several phenomena, trends could be analyzed on the basis of web search streams, and research studies have demonstrated the same. The health care sector, tourism, and stock markets have emerged as three major domains among others.

In the context of financial markets, a prominent impact on stock markets on the basis of large volume and rapidly disseminated data is quite evident. The connection between the market movements and web search-based data cannot be ignored, and this is becoming a significant field of study and research as well. A study of Tantaopas, Padungsaksawasdi & Treepongkaruna (2016), stated that internet search intensity has been emerging as a better direct proxy for investors attention in comparison to earlier indirect proxies used. There are different platforms available which can be utilized to gather the online data for the purpose of predictions. Out of those platforms, Google Search Volume Index (GSVI) is the most utilized one (Agarwal, Kumar & Goel, 2019).

Google Trends: As a Web-Based Searching Tool

Based on the queries posted on Google across various regions and languages for any search term keyword, a platform is provided by Google known as ‘Google Trends’ (Google Search Volume Index), introduced in 2012. Earlier since 2008, Google provided with Google Insights for Search (GIS), a service to display data on search trends which later emerged as Google Trends in 2012. It is considered as a useful tool for determining the attention and sentiments of an investor towards a particular search term. It does not provide an absolute measure, instead a normalized and scaled relative measure is constructed which represents values from 0 to 100, wherein 0 indicates the lowest search volume for that particular term during that duration and 100 indicates

the highest search volume for the given period. This data is usually available on Google Trends on a weekly basis.

The data from here can be gathered for different countries together and can also be confined to a specific country as well. The filters can be applied to categories of the search term (like all categories, business, entertainment, finance, etc.) to remove the overlapping usage of the same term in different fields and further can be extended to channels of search (like youtube, image search, news search, google shopping or web search) to precise the identification of search queries.

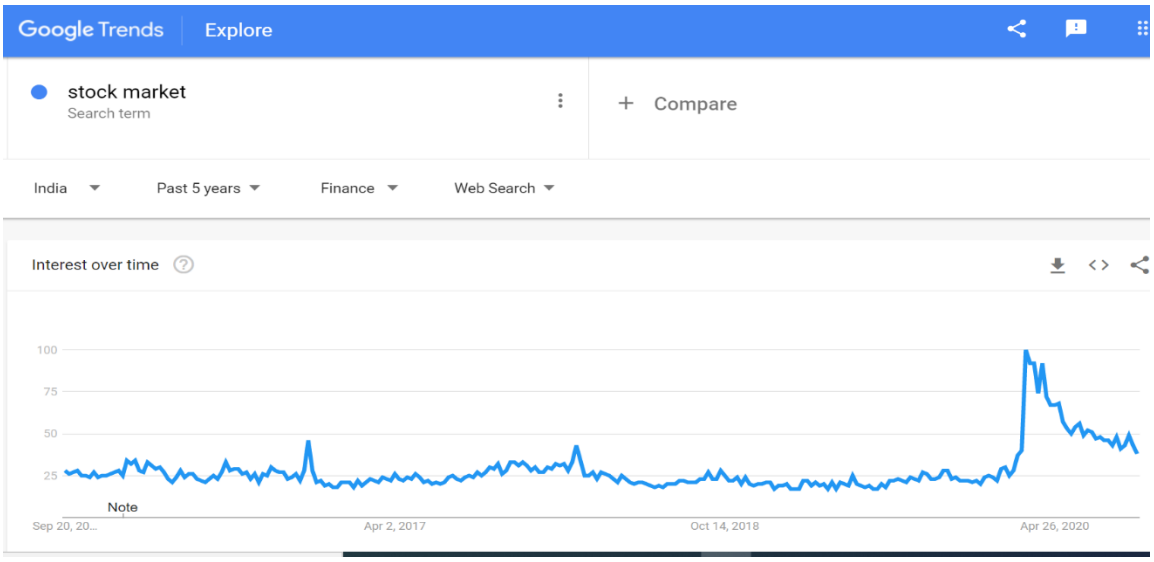

Figure 3 : Google Search Volume Index for Search word “Stock Market” from 2016-2020 (Google Trends,2020)

Google Search as a proxy for Investors’ Sentiments in Financial Markets

This subsection establishes the role of investors behavior with the help of web search-based queries. As there is no standard measure to determine investor’s sentiments, researchers have utilized various methods to gauge the sentiments. Majorly surveys and indexes like Investor Index, Consumer Confidence Index (CCI), FEARS Index etc have been used in the past. The change witnessed in the past few years in technology has given a huge amount of data and information to various fields. In today’s times the web search has emerged as the most important platform to gather information related to the behavior of individuals. A user’s attention or the sentiments could easily be traced with his online search activities. Various economic agents need individual’s behavior and attention related information. The search queries on Google, social media posts submitted across the world by millions of people every day are undoubtedly the good indicators for the same.

With a specific focus in the field of finance, there is always a desire to predict stock market prices. A real problem can be witnessed to model the continuously changing trends in financial markets and to generate meaningful real time predictions about those significant changes. The present ongoing trend portrays the extensive usage of internet and social media platforms to represent the opinion and sentiments about each and every event. This area has emerged as a fascinating field of research as a huge amount of data is available, and a lot of exploitation and extraction is possible to generate eloquent results and new knowledge.

The motivation behind this study is to explore the investor's sentiments represented through their web activities and its influence on stock market movements. The work is interdisciplinary in nature and combines two subject areas, Information Technology and Finance. A lot of new avenues can be explored in this area.

A lot of research has majorly focused on the developed countries and especially United States (US) Markets, but studies are also emerging for developing countries. The studies conducted for developed countries do not have the same implications for developing countries as the financial knowledge is considered to be less in developing countries than that of developed countries, and this difference creates the different level of investor attention and investor sentiment in each country. Across different countries, the irrationality of investors behavior and the degree of inefficiency in financial markets differs. So, generalization of studies across borders is not much possible, and there are very few studies in the context of Indian Stock Markets. This study is to be conducted with a focus on emerging developing countries India and the Indian Stock Markets.

There is strive to gain knowledge that how being one of the largest internet user base country can exploit the information available through internet sources and how such data can be utilized to predict or forecast financial market returns. The active search queries on the internet could be a direct and possible estimate of investor attention. The current study proposes to explore the forecasting power of investors' sentiment on returns. It focuses on determining that how in the light of behavioral finance GSVI can be utilized to forecast stock market returns.

LITERATURE REVIEW

This section comprises the theoretical framework defining the emergence of behavioral finance as a field of research and how web search-based data and social media posts data can be directly proxied for investors' sentiments to generate stock market predictions.

In the 1970s, Efficient Market Theory (EMT) dominated the academic circles. Traditionally, it had been followed that market prices reflect all available information about the market and markets are efficient in nature (Fama, 1976). The research work of Kahneman & Tversky (1979) developed the prospect theory, which explained the decision-making in uncertainty and revealed that during the energy crisis in the 1970s, the results were not consistent to efficient markets theory. The different theoretical models emerging at that time like prospect theory (Kahneman & Tversky, 1979), the concept of Noise (Black, 1986) focused on investor attention and behavior in the stock markets and stock returns and contradicted the fundamentals of EMT. In the 1980s the concept of Behavioral Finance evolved in the academic finance. The concept elaborates that with a contradiction to traditional theories, markets are considered inefficient and irrational investor behavior also plays a major role in predicting market returns or stock prices. Research by Merton (1987) suggested the association of investors attention with stock prices and its liquidity and also raised concerns regarding the difficulty of measurement the degree of investor's attention. The study conducted by Barberis, Shleifer & Vishny (1998) defined investor sentiment as the way investors form beliefs. Research work undertaken by Shiller (2003) explains investor sentiments as a belief about investment risks which cannot be only justified with the facts present. Sentiment is basically the perception of an individual which directly influences his decision-making, and the same is reflected in financial markets. In their study, Kumar & Goyal (2015) stated that the concept of Behavioral Finance incorporates an aid to understand the investor's behavior and challenges the perspective of efficient markets. There is an existence of behavioral biases which deviates the investor from rational investment decisions to irrational ones. Over the years, the role of investors sentiments is justified by research on the theoretical framework. The real question is no longer related to the determination of whether the stock prices are affected by sentiments or not, but rather quantification of its effect and means by which investor sentiments can be measured (Baker & Wurgler, 2007).

Researchers in their studies have utilized different ways to measure the sentiments of investors over time. Broadly two categories of measures have been used: Direct measures as well as Indirect measures. Direct measures include methods related to surveys like Consumer Sentiment

Index (CSI), Investor Intelligence Index (II Index) utilised by (Brown & Cliff, 2004). Indirect measures include indices constructed on the basis of market variables. One of the seminal papers of Baker & Wurgler (2007) introduced the Wurgler Index. These methods had their own limitations in the form of timeliness of the data.

Investors always keep on looking for relatively fast and economical sources of information. Over time, with increasing advent of the internet in human lives, new measures of investor sentiment have positioned itself in the market. It is based on the information that flows through various online internet communication platforms via social media networks and microblogging sites. With the help of Internet Search Engines, the information gathering process has become quite easier, swifter, and inexpensive too. Agarwal et al. (2019) in their literature review paper, revealed that social media platforms which are utilized majorly for research purposes include Internet Search Volume, Social Media Platforms like Facebook and Twitter, and internet message boards. The figure 4 below represents the proportion of choice of the platform and the most sought-after platform is Google Search Volume Index (GSVI).

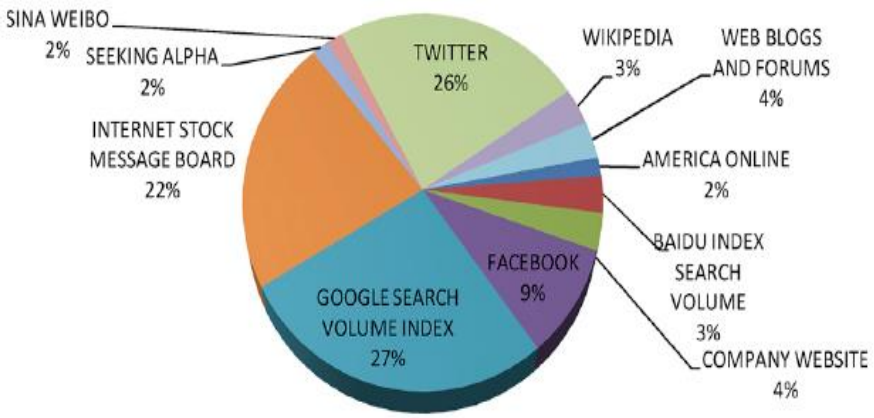

Figure 4 : Proportion of Social Media Platforms Utilized for Research Purposes (Agarwal et al., 2019).

For the very first time, Ginsberg, Mohebbi, Patel, Brammer, Smolinski & Brilliant (2009) identified a new and more precise way to quantify sentiments by gauging them via web search intensity of a keyword which is measured by GSVI (Google Search Volume Index). Research work of Da, Engelberg & Gao (2011) introduced GSVI in the context of investors' attention and stock prices behavior and tested previous traditionally used proxies along with GSVI to predict stock prices. In one of the seminal papers by Joseph, Wintoki & Zhang (2011), the authors proxied online ticker searches for investor sentiments and identified that prediction of abnormal

stock returns and trading volumes is possible with online search intensity. The study revealed that there is existence of more sensitivity of highly volatile stocks to online search activity in comparison to fewer volatile stocks. The focus of several studies by different researchers is to determine the relationship of investor sentiment via their online web search query-based activity and stock trading behavior. The following table represents that how different studies have utilized various online available information at different platforms like GSVI, Twitter, Financial News data and which financial instruments are considered as the representative for the same.

Table 2. Studies Conducted Utilizing Web Search Data

AUTHOR, YEAR	DATA PERIOD	DATA TYPE	COUNTRIES	FINANCIAL INSTRUMENT USED	STATISTICAL TOOLS APPLIED
Da et al. (2011)	2004-2009	GSVI	US	Russell 3000	VAR, Price Pressure Hypothesis
Tantaopas et al. (2016)	2004-2014	GSVI	6 Developed & 4 Developing Countries	Each Country Index	Pairwise Granger Causality, Quantile Regression & VAR
Afkhami, Cormack & Ghoddusi (2017)	2004-2016	GSVI	New York, Brent Crude Data	90 KEY TERMS. Weekly volatility	Granger Casuality, GARCH, VAR
Jin, Wang, Chakraborty, Self, Chen & Ramakrishnan (2017)	2010-2013	News, GSVI, Twitter	Argentina, Brazil, Colombia, Mexico, and Peru. COLOMBIA STOCK INDEX	language modeling, topic clustering and sentiment analysis for news articles, Lasso regression on GSVI and burst detection and event grouping techniques on Twitter data.	Delta Naïve Bayes (DNB)
Tang & Zhu (2017)	2004-2015	GSVI	US	ADRs	Carhart Four Factor Model
Dash & Maitra (2018)	2002-2014	Sentiment Proxies	Indian Stock Markets	Nifty 50, Mid Cap, Small Cap, PCI, AD Ratio, Turnover, VIX	Wavelet method
Igeh & Wijerathne (2018)	2013-2018	GSVI	Norway	Norwegian Mutual Fund	Regression
Vijay, Singh & Malhotra (2018)	2011-2016	Financial News	India	Leading players of the major sectors Like ITC-FMCG, TCS–IT, HDFC-Banking, ONGC-Energy, SunPharma-Pharmaceuticals, Maruti Suzuki-Automobile and SBI-Banking	VADER for Sentiment Analysis (Python) CORRELATION

Asif Khan, Ahmad, Azmi & Akbar, (2019)	2010-2017	GSVI	Islamic stock markets for seven United States (US) Islamic stock indices	118 KEYWORDS FROM FEARS DICTIONARY, Weekly Volatility	Granger Causality, GARCH, new sentiment index, in sample and out sample stock market volatility
Chen & Lo (2019)	2004-2016	GSVI	Taiwan	Taiwan Stock Exchange	Correlation & Regression
Kim, Lučivjanská, Molnár & Villa (2019)	2012-2017	GSVI	Norway	Oslow Stock Exchange	Panel Data Regression with fixed and random effects
Lee (2019)	2016-2018	GSVI	Malaysia	Malaysian Stock Exchange	Regression Analysis
Swamy, Dharani & Takeda (2019)	2012-2017	GSVI	India	S&P BSE 500	Quantile Regression

(Source: Author's Own Compilation)

Literature on Google Search Volume Index and its utility in predicting the financial market returns have utilized different scales and different approaches which has led to varying conclusions. Several studies proved the predictable power of web search-based queries by using different economic and social indicators, while others contradicted the same. The findings of Bijl, Kringhaug, Molnár & Sandviket (2016) and Da et al. (2011) who empirically determined the predictive power of Google search queries were contradicted by Challet & Bel Hadj Ayed (2014) who denied the predictive power of Google Search Volume Index and equated it with the historic returns. Similarly, Kim et al. (2019) resulted in suggesting GSVI carries no information which can produce robust trading signals.

RESEARCH GAPS

The literature is infancy and in the growing stage. The GSVI was first utilised by Ginsberg et al. in (2009) and in context of investors' attention behavior Da et al. (2011) conducted the first study to utilize GSVI as direct proxy for investor attention. This domain of literature is very recent in nature and literature also appraises the informational value of Google Search Volume Index. The present study is an attempt to extend the literature in the context of the utility of web search query-based data to undertake investment decisions in stock markets, especially in the Indian Context. The bibliometric study by Costa, Carvalho & Moreira (2019) determines the extant of growing literature for developed economies and the scope for exploring the developing economies still exists. The studies on other developing nations are also emerging, and a lot of potential is there to explore the Indian stock markets as well. Evidences from literature corroborates the difference in sentiments across different sectors (Baker & Wurgler, 2006; Ekinci & Bulut, 2020), so sectoral analyses is considered to explore the responses of individual sectors. The present study aims at determining the role of data-driven by web-based search queries while taking stock market investment decisions and how it differs along different sectors. The non-availability of universal measure of sentiments allows for varied exploration of the research area. Technological advancement has provided a new line of research and offers a lot of new research scope to obtain improved insights on the relationship of sentiments and market returns. Confusion spotting is also recognized for the purpose of defining the gap as different authors have given contradictory results on the utility of Google Search Volume Index to predict the market returns. Prior studies are limited in number and usually has examined the bivariate linkages between Google Trends and market returns. No such study has focused on the dynamic relationship between Google Trends and stock returns in multivariate set up in Indian context.

DESCRIPTION OF BROAD TOPIC

With the emergence of the concept of behavioral finance, the market efficiency and assumption of investors absorbing all information has been questioned continuously. The research undertaken by De Long, Shleifer, Summers & Waldmann (1990) defines Investor sentiments as the investment risk and cash flows cannot be justified only with the facts at hand. In the early 1980s, when the concept of behavioral finance came into existence, stating the importance of sentiments in financial markets, the sentiments were measured directly via surveys or indirectly via economic and financial variables. These methods had their own limitations of timeliness of data. Surveys do not provide in time information and broadly provides a blend of previous and current opinions. No specific theory has yet conceptualized the irrational behavior of investors and the influence this aspect of an investor exerts on financial markets. The Internet is becoming one of the most popular sources of information. Gradually, online search has become a diurnal routine. Study of Fallows (2005) identified that the experience of people while using internet search engines is positive, and the results obtained by users for their web search queries is trusted by them. Also, the people are very confident of their searching abilities. In order to determine current trends and opinions, web search is the rich source of real time information. The public mood reflected on internet and social media platforms reflect the whole of society on a large scale. The growing interest to employ this previously unavailable online data for predictive purposes has opened doors for research in a wide variety of applications. People from the research community has grabbed a lot of information provided by online platforms to identify the dynamics or predict the trends in various domains of the research. The health care sector, tourism, and stock markets have emerged as three major domains among others which have shown a great potential to utilization of large amount of data available through Google for making certain useful predictions.

The unpredictability of stock markets leads researchers to explore a lot of different avenues to define the stock market trading strategies. There could never be one perfect trading strategy that suits all situations. The studies based on web search-based data or any other social media platform data has limitation of uncertainty about quality and quantity of information, which gives way to further exploration of the stream as well. Various studies have occasioned contradictory results. Work done by Kim et al. (2019) has given contradictory results that web search-based data neither correlate nor predict future abnormal returns. It simply increases volatility and trading volume. This nature of data itself provides an opportunity to explore the data in a lot many ways and opens doors for various avenues of research in the same.

OBJECTIVES OF THE STUDY

The main objective of the study is to examine the association of Indian Stock Market Returns, both the National Stock Exchange (NSE) Nifty and Bombay Stock Exchange (BSE) Sensex stocks market returns and investor's sentiments measured by Google search Volume Index (GSVI). The investor's sentiments are modelled as predictors to forecast stock market returns. The objectives of the study are:

1. To assess whether web search query data measured by Google Search Volume Index can be viewed as a direct proxy for investor's sentiments in relation to NSE Nifty and BSE Sensex stock markets' returns in India.
2. To analyse the reflection of investor's sentiments by Google Search Volume Index for Indian stock market returns across major NSE Nifty and BSE Sensex sectoral indices.
3. To determine that whether the sentiments revealed by Google Search Volume Index embraces forecasting power for the Indian stock market returns.
4. To analyse the significance of investor's optimistic and pessimistic sentiments through Google Search Volume Index in predictability of Indian stock market returns.

RESEARCH METHODOLOGY

RESEARCH QUESTIONS

RQ1: Is there any causal relationship between web search-based query data measured by Google Search Volume Index and the National Stock Exchange (NSE) Nifty and Bombay Stock Exchange (BSE) Sensex stock markets' returns?

RQ2: Are NSE Nifty and BSE Sensex stock markets' returns aligned to investor's sentiments displayed by web search behavior of individual investors?

RQ3: Is there any difference in investor's sentiment in relation to stock market returns across major NSE Nifty and BSE Sensex sectoral indices?

RQ4: Can Indian stock market returns predictability be improved with the help of investor's sentiments measured through Google Search Volume Index?

RQ5: What is the role of investor's attention and sentiment in improving market efficiency?

RQ6: What is the nature of relationship of risk association with optimistic and pessimistic market sentiments revealed through Google Search Volume Index?

Hypotheses

Ho: Web Search based investor's sentiments measured by Google Search Volume Index does not affect NSE Nifty and BSE Sensex stock markets' returns movement.

Ho: There is no difference in investor's sentiments with respect to stock market returns across major NSE Nifty and BSE Sensex sectoral indices.

Ho: There is no predictive power of the investor's sentiments revealed by Google Search Volume Index for Indian Stock Market Returns.

Ho: There is no correlation between optimistic and pessimistic search terms with aggregate Indian stock market returns.

PROPOSED RESEARCH METHODOLOGY

Research Design : The proposed study is an exploratory study which reconnoitres the relationship of web search-based query data as a direct proxy for investor's sentiment in Indian stock market returns and further explores its effect on forecasting the Indian stock market returns. The scope of the study defines that first of all whether the web search-based query data on internet measured by Google Search Volume Index (GSVI) can be taken as a direct proxy for investor's sentiment in relation to the Indian stock market returns both at the aggregate market returns level and at the sectoral returns level. Further, can predictability of market returns be improved using the same. The study explores the impact of both optimistic(positive) and pessimistic(negative) measures of sentiments on Indian stock market returns momentum and its predictability of future stock market returns.

Data Collection: The data is proposed to be collected from Secondary Sources. Data is to be collected for Financial time series and keywords identified as investor's sentiment proxy. The financial data for stock market returns and sectoral returns is to be collected from the NSE website. The data regarding the keywords identified for google search is to be collected from the Google Trends website. On google trends for longer periods more than a quarter, weekly data is available. So, it is proposed to collect weekly data for the duration of 10 years beginning from 1st January 2011 to 31st December 2020. The keywords for google search are proposed to be identified from the Harvard IV-4 Dictionary, the Laswell dictionary of General Inquirer. These are the financial dictionaries which are used for the financial keyword identification. It is proposed to prepare two separate constructs for keywords identified. One for positive sentiment and another one for negative sentiment.

Statistical Tools Used: It is proposed to apply tools like Granger Causality, Quantile Regression Models and further calculating MAPE (Mean Absolute Percentage Error) to check the predictive accuracy of a forecasting method. A causality relationship is to identify whether investor sentiments are derived by market returns, volatility or vice versa. Quantile regression models help in defining asymmetric relationships between positive and negative market conditions and attention.

Table 3. Proposed Research Methodology

Research Design	Exploratory Study
Scope of Study	Stock Market Returns Predictability
Data	Financial Time Series Data and Google Trends Data
Source	Secondary Sources
Duration of Data	10 Years beginning from 1st January 2011 to 31st December 2020.
Nature of Data	Weekly data for Nifty 50 and Sensex 30 and other Nifty and Sensex Sectoral Indices (NSE & BSE Website) Weekly data of Key Words identified for Google Search. (GoogleTrends)
Key Word Identification	Financial Dictionaries (Harvard IV-4 Dictionary , Laswell dictionary of General Inquirer).
Constructs of Sentiments from GSVI	2 Constructs: One for Positive Sentiment and One for Negative Sentiment
Statistical Tools to be Used	Granger Causality, Quantile Regression Models, MAPE

(Source: Author's Own Compilation)

PROPOSED EXPECTED OUTCOME

The study proposes to reveal the role of investor's sentiments through their web search-based activity in the stock market returns. The contribution of this strand of literature is very limited in the developing countries. The studies conducted for developing countries is comparatively low in number and a lot of scope to explore Indian Context is possible. The study helps in defining the role of investors sentiments to improve the predictability of the markets. This is new stream of research in the context of big data analytics for sentiment. The spontaneity of behavior has its own interesting signalling properties and the same can be utilized for improved accuracy of short-term and long-term predictions for market returns. The incorporation of internet search queries in existing benchmark forecasting models can benefit the policymakers to improve accuracy in their forecasting models and can help investors in construction of profitable trading strategies.

REFERENCES

Afkhami, M., Cormack, L., & Ghoddusi, H. (2017). Google search keywords that best predict energy price volatility. *Energy Economics*, 67, 17–27.

Agarwal, S., Kumar, S., & Goel, U. (2019). Stock market response to information diffusion through internet sources: A literature review. *International Journal of Information Management, 45*, 118–131. https://doi.org/10.1016/j.ijinfomgt.2018.11.002

Asif Khan, M., Ahmad, R., Azmi, A., & Akbar, M. (2019). A new sentiment index for the Islamic stock market. *Investment Analysts Journal,* 48(2), 146-172.

Barberis, N., Shleifer, A., & Vishny, R . (1998). A model of investor sentiment. *Journal of Financial Economics*, 49(3), 307–343.

Baker, M., & Wurgler, J. (2006). Investor sentiment and cross-section of stock returns. *Journal of Finance,* 61, 1645-1680.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market, *Journal of Economic Perspectives,* 21, 129–151.

Bijl, L., Kringhaug, G., Molnár, P., & Sandvik, E. (2016). Google searches and stock returns. *International Review of Financial Analysis*, 45, 150–156.

Black, F. (1986). Noise. *The journal of finance*, 41(3), 528-543.

Brown, G.W., & Cliff, M.T. (2004). Investor Sentiment and the Near-Term Stock Market. *Journal of Empirical Finance*, 11, 1-27.

Challet, D., & Bel Hadj Ayed, A. (2014). Do Google Trend Data Contain More Predictability than Price Returns? *SSRN Electronic Journal*. Published. https://doi.org/10.2139/ssrn.2405804

Chen, H. Y., & Lo, T. C. (2019). Online search activities and investor attention on financial markets. *Asia Pacific Management Review*, 24(1), 21-26.

Costa, D. F., Carvalho, F. D. M., & Moreira, B. C. D. M. (2019). Behavioral economics and behavioral finance: A bibliometric analysis of the scientific fields. *Journal of Economic Surveys,* 33(1), 3-24.

Da, Z., Engelberg, J., & Gao, P. (2011). The Sum of All FEARS: Investor Sentiment and Asset Prices. *SSRN Electronic Journal*. Published. https://doi.org/10.2139/ssrn.1509162

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, 66(5), 1461-1499.

Dash, S. R., & Maitra, D. (2018). Does sentiment matter for stock returns? Evidence from Indian stock market using wavelet approach. *Finance Research Letters*, 26, 32-39.

De Long, J., Shleifer, A., Summers, L., & Waldmann, R. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, 98(4), 703–738.

Ekinci, C., & Bulut, A. E. (2021). Google search and stock returns: A study on BIST 100 stocks. *Global Finance Journal*, *47*, 100518. https://doi.org/10.1016/j.gfj.2020.100518

Fallows, D. (2005). Search engine users. Pew Internet & American Life Project. Retrieved from http://www.pewinternet.org/pdfs/PIP_Searchengine_users.pdf

Fama, E.F. (1976). *Foundations of finance: Portfolio decisions and securities Prices*. Blackwell.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457(7232), 1012-1014.

Harford, T. (2017). Just google it: The student project that changed the world. *Accessible online http://www. BBC. com/news/business*, *39129619*.

Igeh, S. A., & Wijerathne, I. (2018). *Empirical investigation of the relationship between Google search volume index and mutual fund performance and flows: evidence from Norway* (Master's thesis, University of Stavanger, Norway).

Jin, F., Wang, W., Chakraborty, P., Self, N., Chen, F., & Ramakrishnan, N. (2017). Tracking multiple social media for stock market event prediction. *In Industrial conference on data mining*,10357,16-30. Springer, Cham. https://doi.org/10.1007/978-3-319-62701-4_2

Joseph, K., Wintoki, M. B., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting*, 27(4), 1116-1127.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263-291. doi:10.2307/1914185

Keelery, S. (2020). *Number of internet users in India from 2015 to 2020 with a forecast until 2025(in millions)*. Digital Market Outlook https://www.statista.com/statistics/255146/number-of-internet-users-in-india/

Kim, N., Lučivjanská, K., Molnár, P., & Villa, R. (2019). Google searches and stock market activity: Evidence from Norway. *Finance Research Letters*, 28, 208–220. https://doi.org/10.1016/j.frl.2018.05.003

Kumar, S., & Goyal, N. (2015). Behavioural biases in investment decision making – a systematic literature review. *Qualitative Research in Financial Markets*, 7(1), 88–108. https://doi.org/10.1108/qrfm-07-2014-0022

Lee, K. Y. M. (2019). Google Search and Stock Market Performance: Evidence from Malaysia. *UNIMAS Review of Accounting and Finance*, 3(1), 49-55

Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, 42(3), 483–510. https://doi.org/10.1111/j.1540-6261.1987.tb04565.x

Paule-Vianez, J., Gómez-Martínez, R., & Prado-Román, C. (2020). A bibliometric analysis of behavioural finance with mapping analysis tools. *European Research on Management and Business Economics*, 26(2), 71–77. https://doi.org/10.1016/j.iedeen.2020.01.001

Shiller, R.J. (2003). From efficient markets theory to behavioral finance. *The Journal of Economic Perspectives*, 17(1): 83–104.

Swamy, V., Dharani, M., & Takeda, F. (2019). Investor attention and Google Search Volume Index: Evidence from an emerging market using quantile regression analysis. *Research in International Business and Finance*, 50, 1-17.

Tang, W., & Zhu, L. (2017). How security prices respond to a surge in investor attention: Evidence from Google Search of ADRs. *Global Finance Journal*, 33, 38-50.

Tantaopas, P., Padungsaksawasdi, C., & Treepongkaruna, S. (2016). Attention effect via internet search intensity in Asia-Pacific stock markets. *Pacific-Basin Finance Journal*, 38, 107-124.

Trends.google.com (2020). Google Trends. *Retrieved from https://trends.google.com/trends/explore?cat=7&date=today%205-y&geo=IN&q=stock%20market*.

Vijay, N., Singh, S., & Malhotra, G. (2018). Sentiment Analysis: Gauging the Effect of News on Stock Prices in Indian Stock Market. *International Journal of Trade, Economics, and Finance*, 9(4).

Wang, X., Ye, Q., Zhao, F., & Kou, Y. (2018). Investor sentiment and the Chinese index futures market: Evidence from the internet search. *Journal of Futures Markets*, 38(4), 468-477.

MANAV RACHNA
Multidisciplinary

Dr. O.P. Bhalla Central Library
Manav Rachna International Institute of Research and Studies
(A deemed-to-be-university under section 3 of UGC Act 1956)

SIMILARITY VERIFICATION CERTIFICATE

Title: An Empirical Investigation of Google Trends for Assessing Investor's Sentiments and Predictability of Indian Stock Market Returns.

Researcher: Ms. Divya Jain

Registration Number: 19/Ph.D/05

Supervisor: Dr. Megha Chhabra

Department: Management (FMS)

This is to certify that the above mentioned Short ✓ Synopsis/ ~~Thesis~~ was scanned for similarity index through Turnitin Software on 5/10/2021 with submission ID 1665799365 and the similarity index was found to be 7 %. The complete report running into 26 Pages, 6001 Word count with File Size 3338 KB is being forwarded to the Research Supervisor/DRC/DRB for review.

Dr. O.P. BHALLA CENTRAL LIBRARY * MRIIRS, Faridabad * 6/10/21

Date: 6/10/2021

The complete report of the above Short ✓ Synopsis/ ~~Thesis~~ has been reviewed and the similarity index is below the accepted norms as per the University Grants Commission (Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions) Regulations 2018. The Short ✓ Synopsis/ ~~Thesis~~ may be considered for further processing

Research Scholar

Jain
6/10/2021

Research Supervisor

6/10/21

Short Synopsis DIVYA JAIN 5.10.2021

ORIGINALITY REPORT

7% SIMILARITY INDEX

Excluding matches less than 14 words

3%	6%	0%
INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

plag 6/10/21

PRIMARY SOURCES

1	Paritosh Chandra Sinha. "Noise of Investors' Attention Mania in the Twenty-first-Century Indian Stock Markets: ARDL and Augmented GARCH-X Models", Global Business Review, 2021 Publication	1
2	Shweta Agarwal, Shailendra Kumar, Utkarsh Goel. "Stock market response to information diffusion through internet sources: A literature review", International Journal of Information Management, 2019 Publication	1
3	www.ijtef.org Internet Source	1
4	Hong-Yi Chen, Te-Chien Lo. "Online search activities and investor attention on financial markets", Asia Pacific Management Review, 2019 Publication	<1
5	hdl.handle.net Internet Source	<1

6	Fang Jin, Wei Wang, Prithwish Chakraborty, Nathan Self, Feng Chen, Naren Ramakrishnan. "Chapter 2 Tracking Multiple Social Media for Stock Market Event Prediction", Springer Science and Business Media LLC, 2017 Publication	<
7	lutpub.lut.fi Internet Source	<
8	Muhammad Asif Khan, Rubi Ahmad, Anna Azmi, Muhammad Akbar. "A new sentiment index for the Islamic stock market", Investment Analysts Journal, 2019 Publication	<
9	Parkpoom Tantaopas, Chaiyuth Padungsaksawasdi, Sirimon Treepongkaruna. "Attention effect via internet search intensity in Asia-Pacific stock markets", Pacific-Basin Finance Journal, 2016 Publication	<1
10	arno.uvt.nl Internet Source	<1
11	Neri Kim, Katarína Lučivjanská, Peter Molnár, Roviel Villa. "Google searches and stock market activity: evidence from Norway", Finance Research Letters, 2018 Publication	<1
12	dergipark.org.tr Internet Source	

		<1 %
3	Malcolm Baker, Jeffrey Wurgler. "Investor Sentiment in the Stock Market", Journal of Economic Perspectives, 2007 Publication	<1 %
4	www.sciencegate.app Internet Source	<1 %
5	Submitted to Caleb University Student Paper	<1 %
6	Canh Phuc Nguyen, Christophe Schinckus, Thai Vu Hong Nguyen. "Google search and stock returns in emerging markets", Borsa Istanbul Review, 2019 Publication	<1 %
7	Vighneswara Swamy, M. Dharani, Fumiko Takeda. "Investor attention and Google Search Volume Index: Evidence from an emerging market using quantile regression analysis", Research in International Business and Finance, 2019 Publication	<1 %

Exclude quotes On

Exclude bibliography On

Exclude matches < 14 words